Mathematically Reduced Chemical Reaction Mechanism Using Neural Networks

N. K. Butuk

K. Johnson

Prairie View A & M University

OVERVIEW

- **SIMPLIFICATION OF HIGHER-DIMENSION DYNAMICAL SYSTEMS VIA NPCA NEURAL NETWORKS**

- **APPLICATION TO REACTIVE FLOWS (COMBUSTION)**

COMPETING TECHNIQUES

- STEADY STATE APPROXIMATION
- PARAMETER LUMPING
- ILDM-CPS
- SINGULAR PERTURBATION THEORY
- CENTER MANIFOLD REDUCTION

Objective

- **Implement Non-Linear PCA (NPCA)**
- **PCA-Principal Component Analysis**
- **NPCA Implemented as Neural Network**
- **Speed Up Training of NPCA via**
- **Techniques of Kernel Smoothing**

OUTLINE

- **INTRODUCTION-Neural Network Basics**
- **NONLINEAR PRINCIPAL COMPONENT ANALYSIS**
- **DETAILED MATHEMATICAL ANALYSIS**
- **RESULTS:**
- **. IMPLEMENTATION: SAMPLE MECHANISM)**
- **BRIEF DESCRIPTION OF FLOW SOLVER – KEN JOHNSON**
- **CONCLUSIONS**

Neural Network Basics..1

- **Artificial NN Consists of Computational Units Called Neurons**
- **Receive a Number of Input Signals**
- **Produce an Output Signal**
- **Real Valued Functions on R^n**

$$h(x) = \sigma(w^T x + b)$$

$$x \in \mathbb{R}^n$$

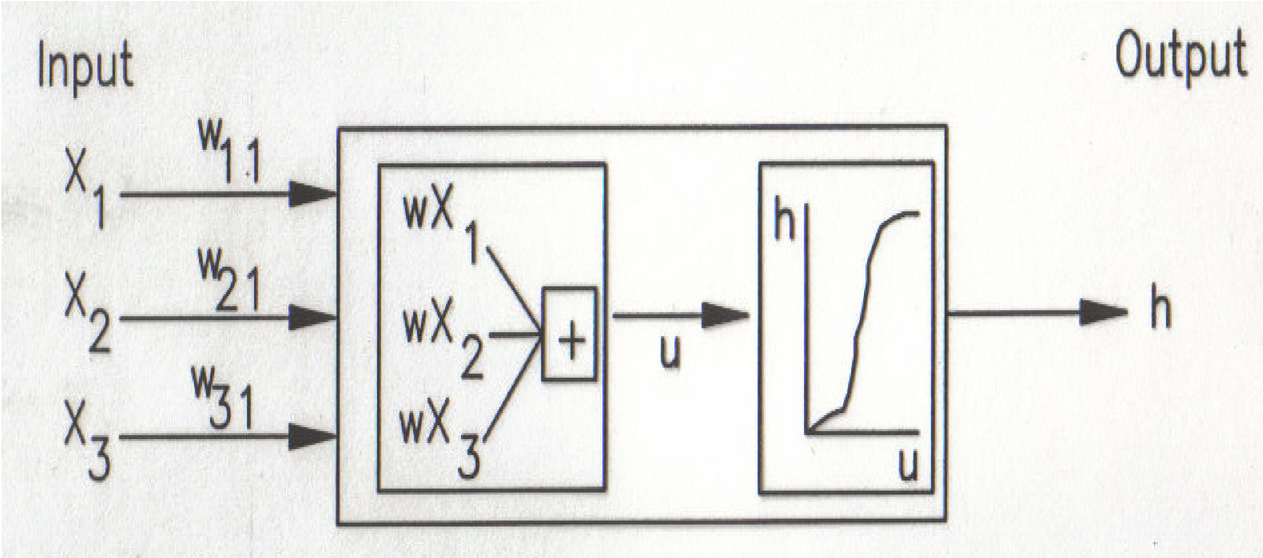
$$w^T = (w_1, w_2 \ldots w_n) \in \mathbb{R}^n$$

$$b \in \mathbb{R}$$

$$\sigma(x) - is - known - as - sigmoid - function$$

Neural Network Basics..2

- **Neuron**



Schematic of a Neuron

Neural Networks Basics…3

- **Single Neuron (also known as Perceptron) Can only estimate linear functions.**
- **For General Non-Linear Problems MultiLayer Perceptrons (MLP-NN) are usually used.**
- **Feedforward Neural Network with one Hidden Layer Containing k neurons approximates a Real Valued Non-Linear Function on R^n as**

$$m(x) = \sum_{i=1}^{k} c_i \sigma(w_i^T x + b_i) + c_o$$

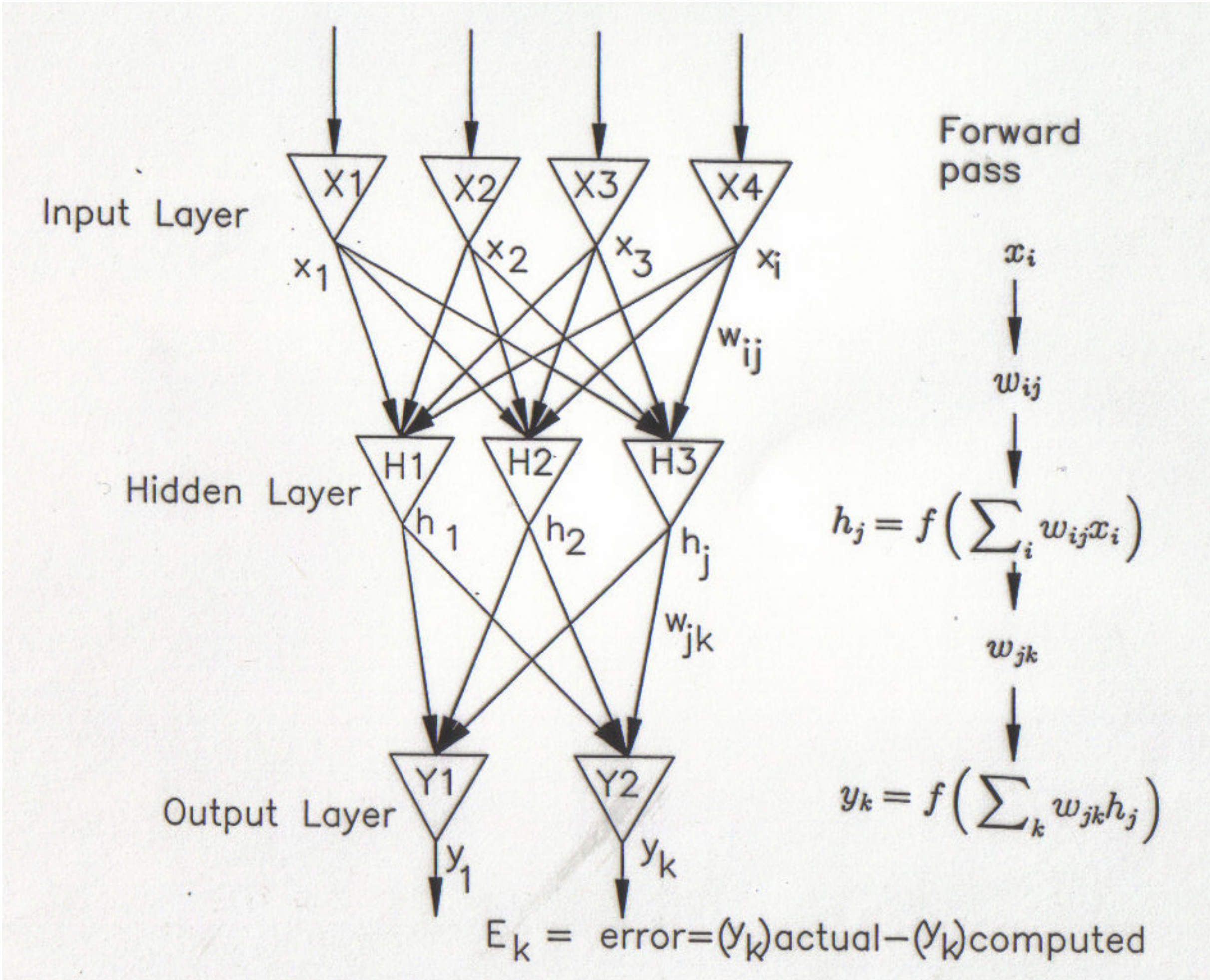
$$symbols - as - above - with$$

Neural Networks Basics…4

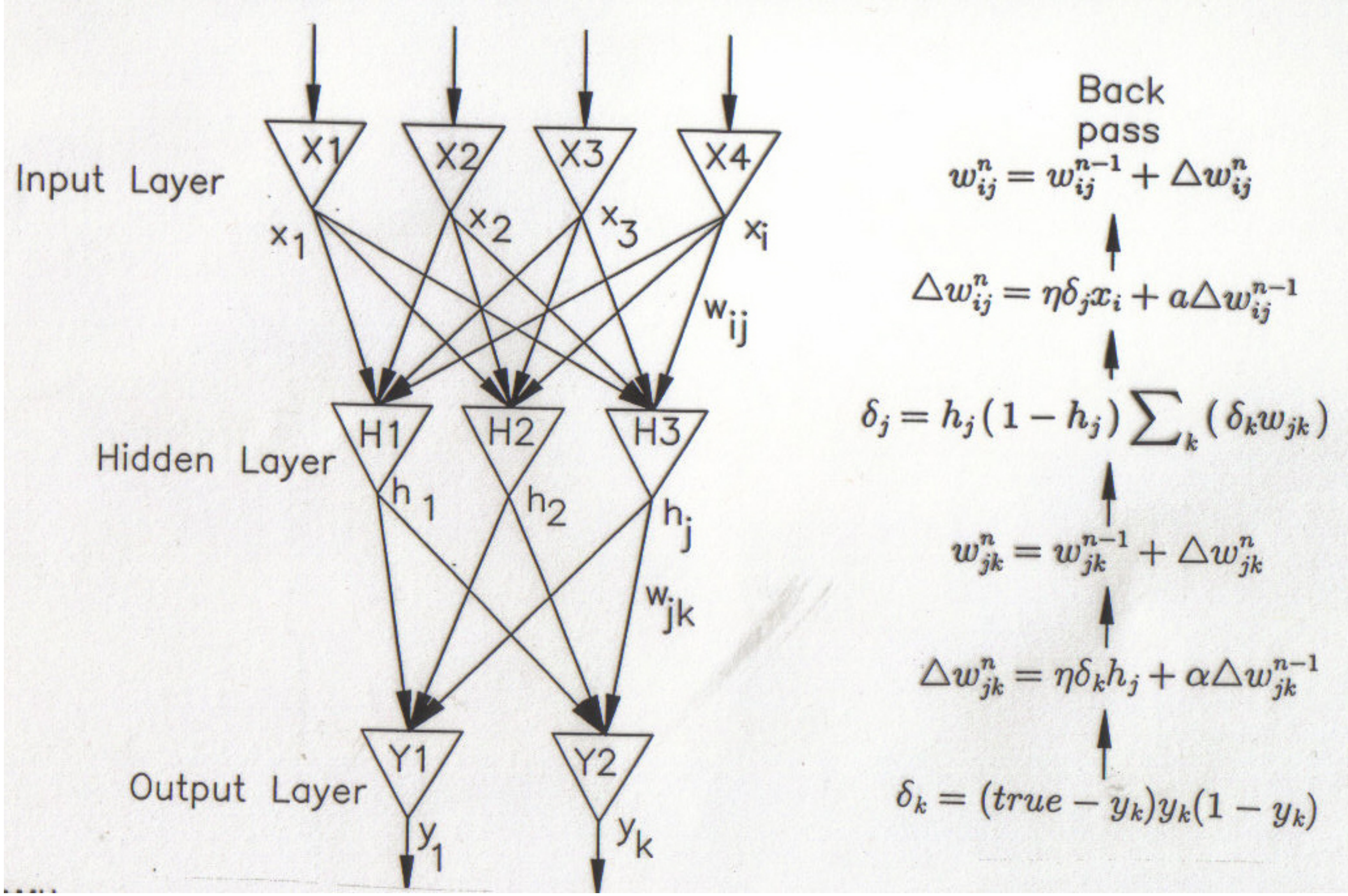
$$b_1 \ldots b_k, c_o, c_1 \ldots c_k \in \text{¡}$$

- **w’s,b’s,c’s are parameters that specify a NN**

Neural Networks Basics…5



Neural Networks Basics…6



Neural Networks Basics…7

- **Given (iid) Training Data, Chosen at Random from a population (X,Y) {(X_1,Y_1)…(X_m ,Y_m)}**
- **The Parameters of the Network are chosen to minimize the empirical L_2 risk**

$$\frac{1}{m}\sum_{j=1}^{m}\left|m(X_j)-Y_j\right|^2$$

Neural Networks Basics…8

- **There are No Practical Algorithmic for finding the Global Minimum of L_2**
- **Algorithm Described above is the Popular BackPropagation**
- **Back Propagation (using steepest descent) often converges to a local minimum**

Nonlinear Principal Component Analysis (NPCA)

- **1. Need for Mechanism Reduction**
  - **Typically can consists of >500 species**
  - **>2000 elementary chemical reactions**
- **2. Full chemistry model will have to solve**
  - **>500 coupled Differential Equations**
  - **Computationally prohibitive**

NPCA…2.

- 3. Most reaction state space maybe redundant
  - Active space may exist as manifolds of lower dimension
- 4. NPCA global transformation mapping from
  - High dimensional state space (n) to
  - Low dimensional manifold (m)

NPCA…3

- **5. Let:**

$$X \in \mathbb{R}^n$$

$$Y \in \mathbb{R}^m$$

$$m < n$$

$$X = f(Y)$$

- **Intrinsic Dimensionality of Data really m not n**

NPCA-NN…4

$$G : X \rightarrow Y$$

$$H : Y \rightarrow X$$

$$H \circ G : X \rightarrow X$$

- **Need sufficient x-data points in state space - a number of representative trajectories**

Generating Trajectory Data

- **Need representative data within flammability limits of a given fuel**

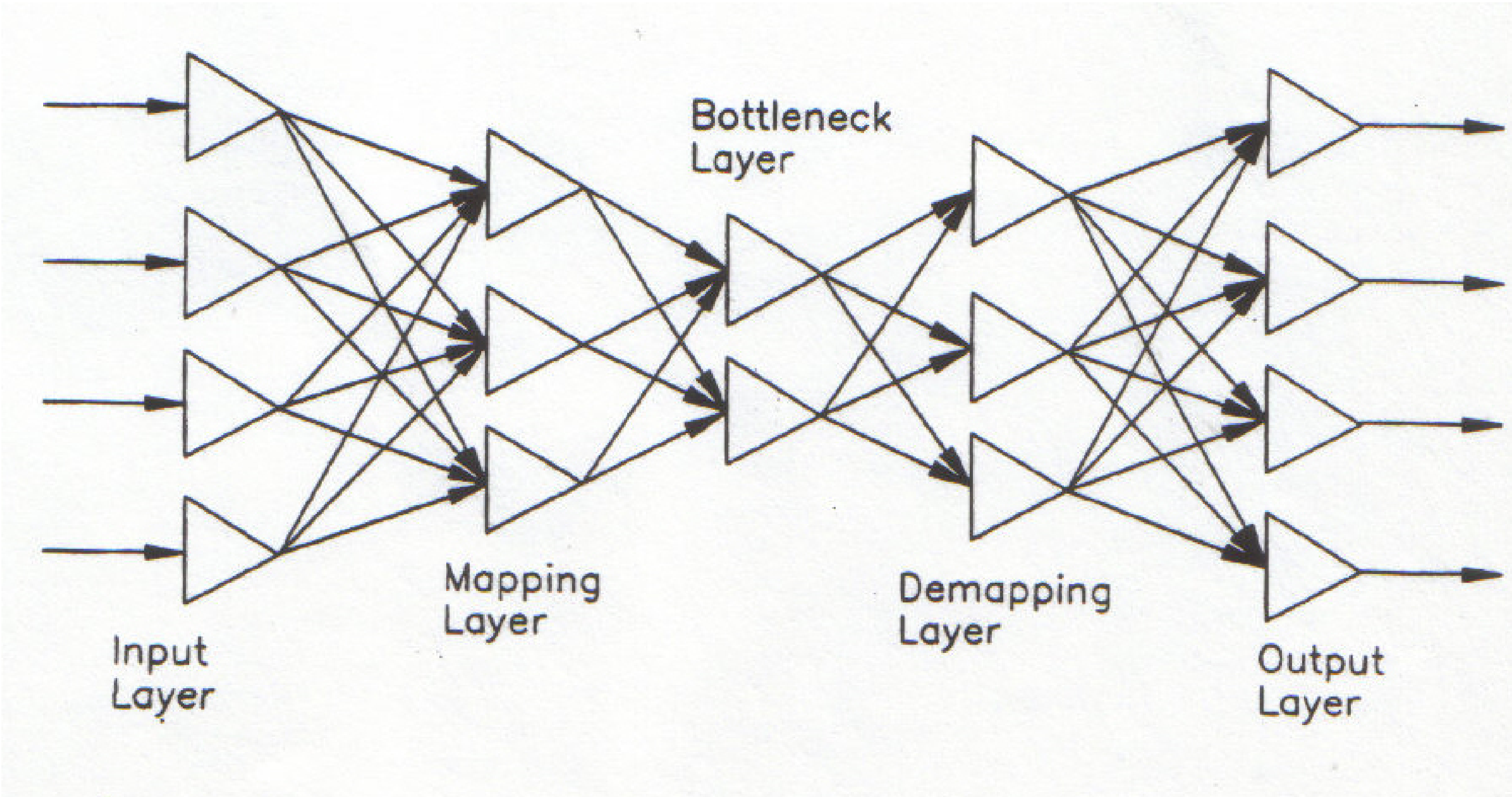
- **Can use D.O.E. Software**

- **Trajectories Generated at Random**

NPCA-NN Algorithm

- **Consists of:**
  - **Two standard Multilayered Perceptrons (MLP)**
  - **First network implements G**
  - **Second network implements H**
  - **The "Bottleneck Layer" is the reduced dimension**

NPCA-NN Schematic



Training in stages

- **Accelerates convergence by several orders of magnitudes**

- **Problem: Combining Networks is a non-linear operation**

- **Several methods of combining available - still active area of research**

Results

- **Test Mechanisms**
  - **Bromide acid synthesis**

Bromide acid synthesis Mechanism

$$Br_2 \xrightarrow{k_1} 2Br$$
$$2Br \xrightarrow{k_2} Br_2$$

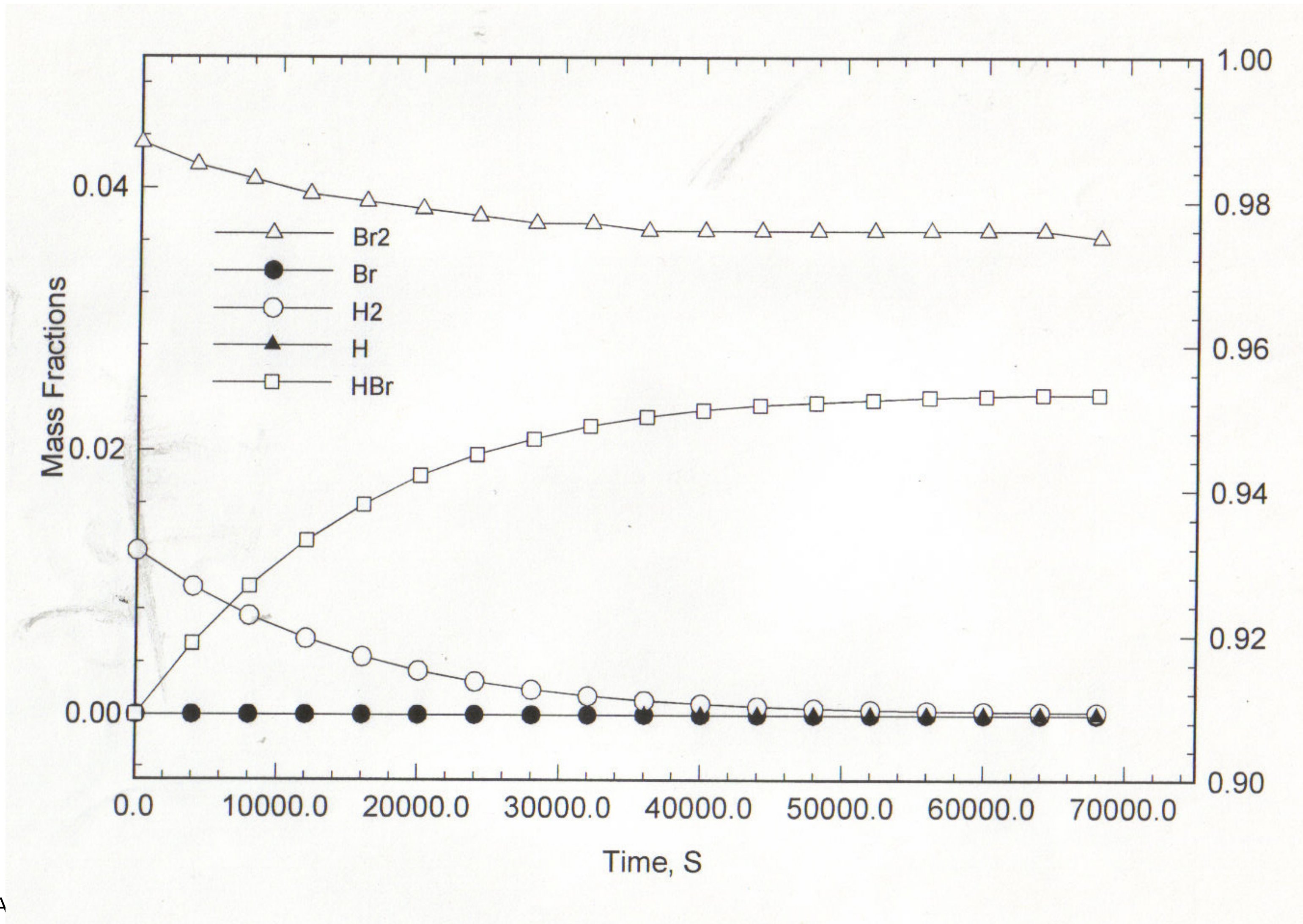
$$H_2 \xrightarrow{k_3} 2H$$
$$2H \xrightarrow{k_4} H_2$$

$$H + Br \xrightarrow{k_5} HBr$$
$$HBr \xrightarrow{k_6} H + Br$$

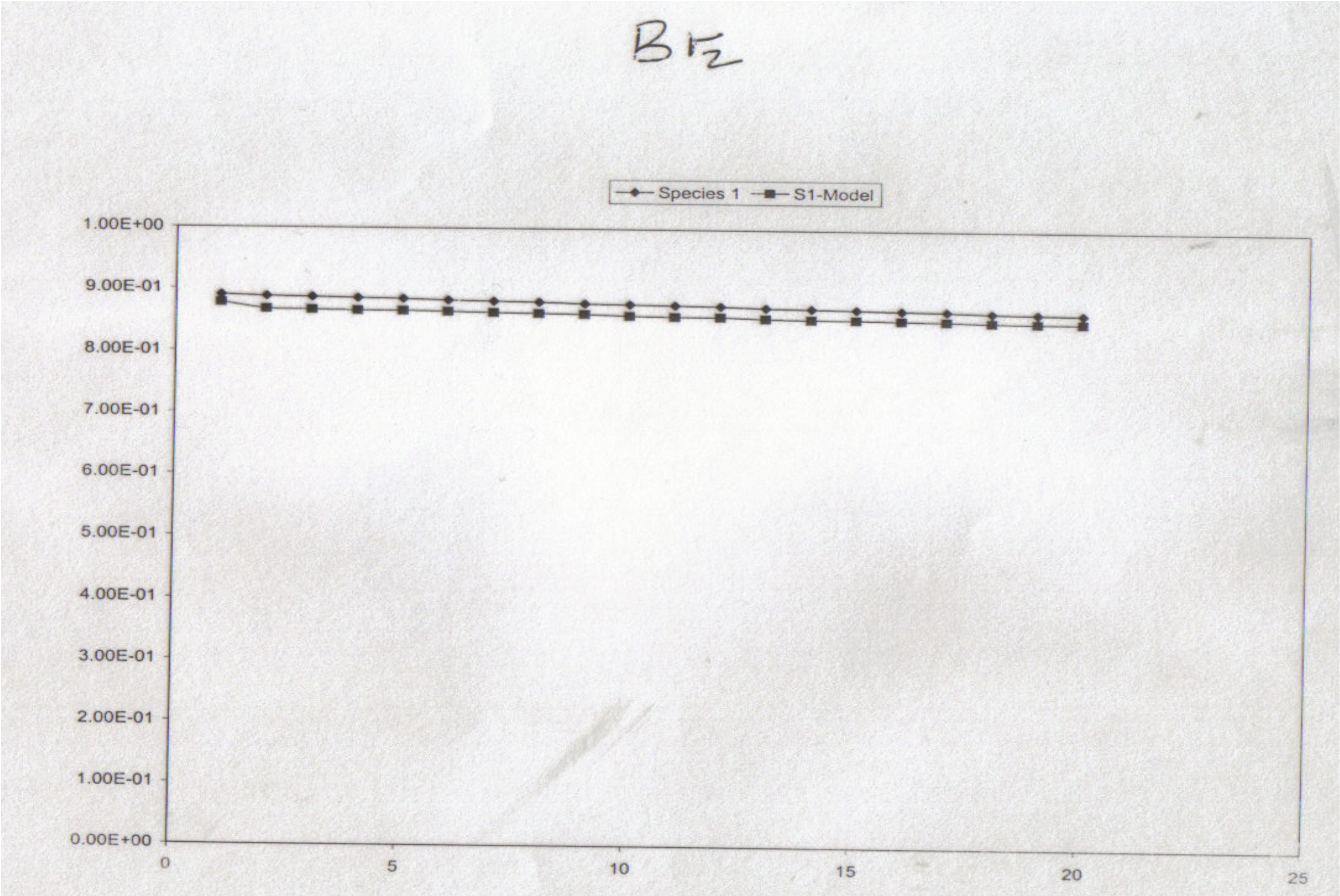
$$\frac{d\,[Br_2]}{dt} = k_2\,[Br]^2 - k_1\,[Br_2]$$
$$\frac{d\,[Br]}{dt} = 2k_1\,[Br_2] - 2k_2\,[Br]^2 + k_6\,[HBr] - k_5\,[H]\,[Br]$$
$$\frac{d\,[H_2]}{dt} = k_4\,[H]^2 - k_3\,[H_2]$$
$$\frac{d\,[H]}{dt} = 2k_3\,[H_2] - 2k_4\,[H]^2 + k_6\,[HBr] - k_5\,[H]\,[Br]$$
$$\frac{d\,[HBr]}{dt} = k_5\,[H]\,[Br] - k_6\,[HBr]$$

$$k_1 = 9.2E-5, \quad k_2 = 4.0E15, \quad k_3 = 9.2E-5$$
$$k_4 = 4.0E15, \quad k_5 = 1.0E15, \quad k_6 = 1.0E-5$$

Results-1

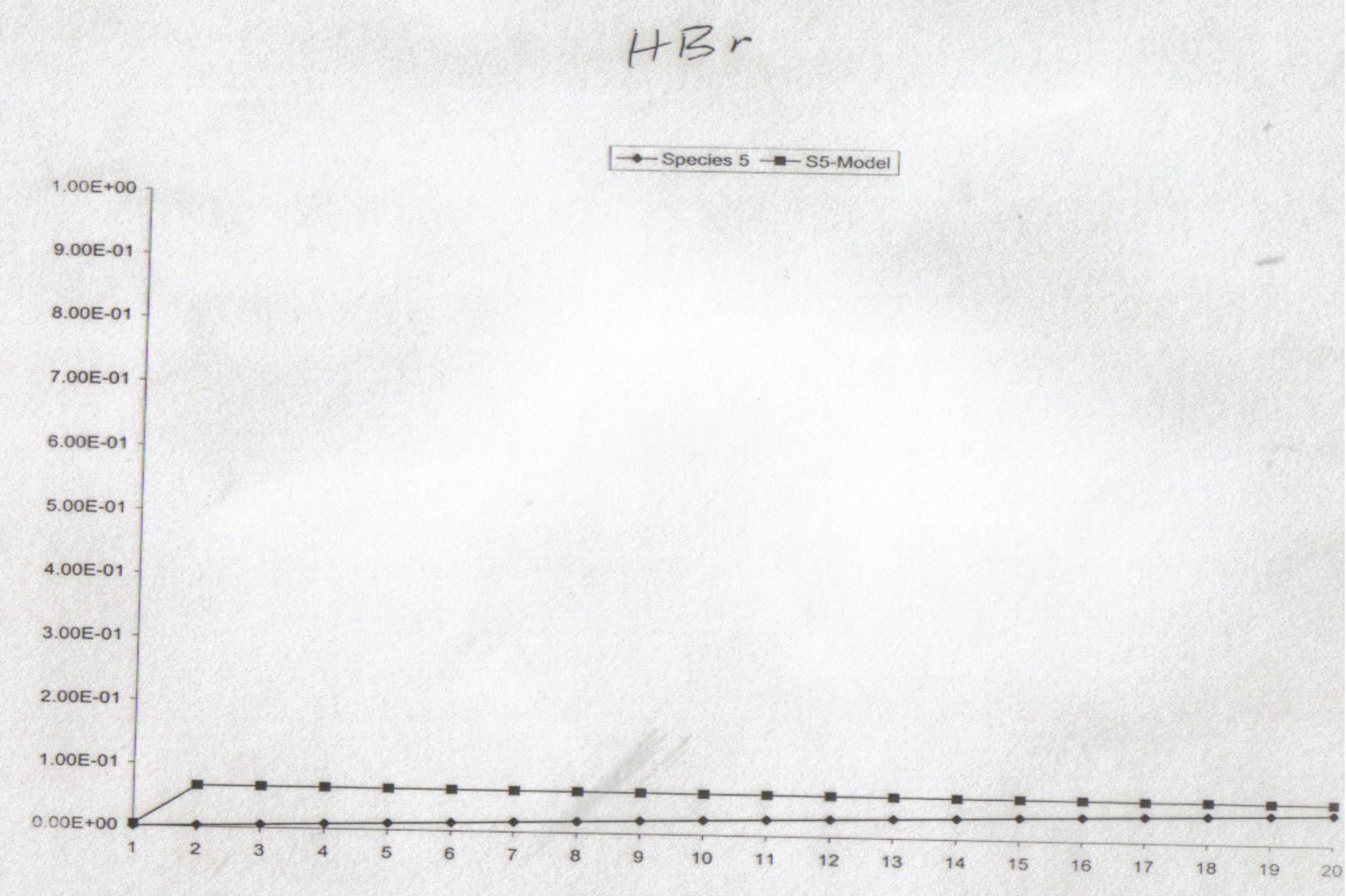

Br2
Br
H2
H
HBr
Mass Fractions
0.04
0.02
0.00
1.00
0.98
0.96
0.94
0.92
0.90
0.0
10000.0
20000.0
30000.0
40000.0
50000.0
60000.0
70000.0
Time, S

Results-2



Br2
Species 1
S1-Model
1.00E+00
9.00E-01
8.00E-01
7.00E-01
6.00E-01
5.00E-01
4.00E-01
3.00E-01
2.00E-01
1.00E-01
0.00E+00
0
5
10
15
20
25

Results-3


HBr
Species 5
S5-Model
1.00E+00
9.00E-01
8.00E-01
7.00E-01
6.00E-01
5.00E-01
4.00E-01
3.00E-01
2.00E-01
1.00E-01
0.00E+00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Function Approximation…1

- We have a population consisting of random vectors (X,Y). X is R^n – valued and y is R-valued
- Looking for a function
- $f: R^n \to R$
- Our Criteria for finding f is that the expectation of L_2 norm (risk) is small
- Note $L_2 = |f(X)-Y|^2$, is also random
- Let m(x) be the function that minimizes $E\,|f(X)-Y|^2$
- Function m(x) is unknown, we only have sampled data to estimate it.

Function Approximation…2

- **Given a function g(x) of a random variable x**

$$\mathrm{E}[g(x)] = \int g(x) P(x) dx$$

Where P(x) is the distribution or density of x

Function Approximation…3

Let $\hat{m}(x)$ **be an approximation of m(x) using data**

$$E\left[\left|\hat{m}(x)-Y\right|^2\right]=\int_{\text{¡}^n}\left|\hat{m}(x)-m(x)\right|^2 P(x)dx+E\left|m(x)-Y\right|^2$$

$\hat{m}(x)$ **= m(x) if the first term on right, L_2=0**

Function Approximation…4

- **The important term P(x) is the distribution, or density of X. It is usually also estimated from data.**
- **A popular approach is termed Kernel Smoothing.**
- **Define a Kernel function** $K(\frac{x-X_i}{h})$ **having compact support and symmetric about origin**
- **Also** $$\int K(x)d(x)=1$$

Function Approximation…5

The Density can be approximated from data

$$\hat{P}(x) = \frac{1}{mh}\sum_{i=1}^{m} K\left(\frac{x - X_i}{h}\right)$$

Function Approximation…6

- Now how do we Estimate m(x) from data
- Several different approaches have been developed
  - Parametric
  - Non-Parametric Approaches
- Examples of Non-Parametric methods
  - Neural Networks(NN)
  - Radical Basics Functions Networks
  - Orthogonal series methods (including wavelets)
  - Least squares Estimates using splines
  - Local Polynomial Kernel Estimates
- Our focus is on NN approaches

Function Approximation…7

- **From statistical theories of regression m(x) can be defined as** $m(x) = \mathrm{E}(Y \mid X = x) = \int yP(y \mid x)dy$

$$= \frac{\int yP(x,y)dy}{\int P(x,y)dy}$$

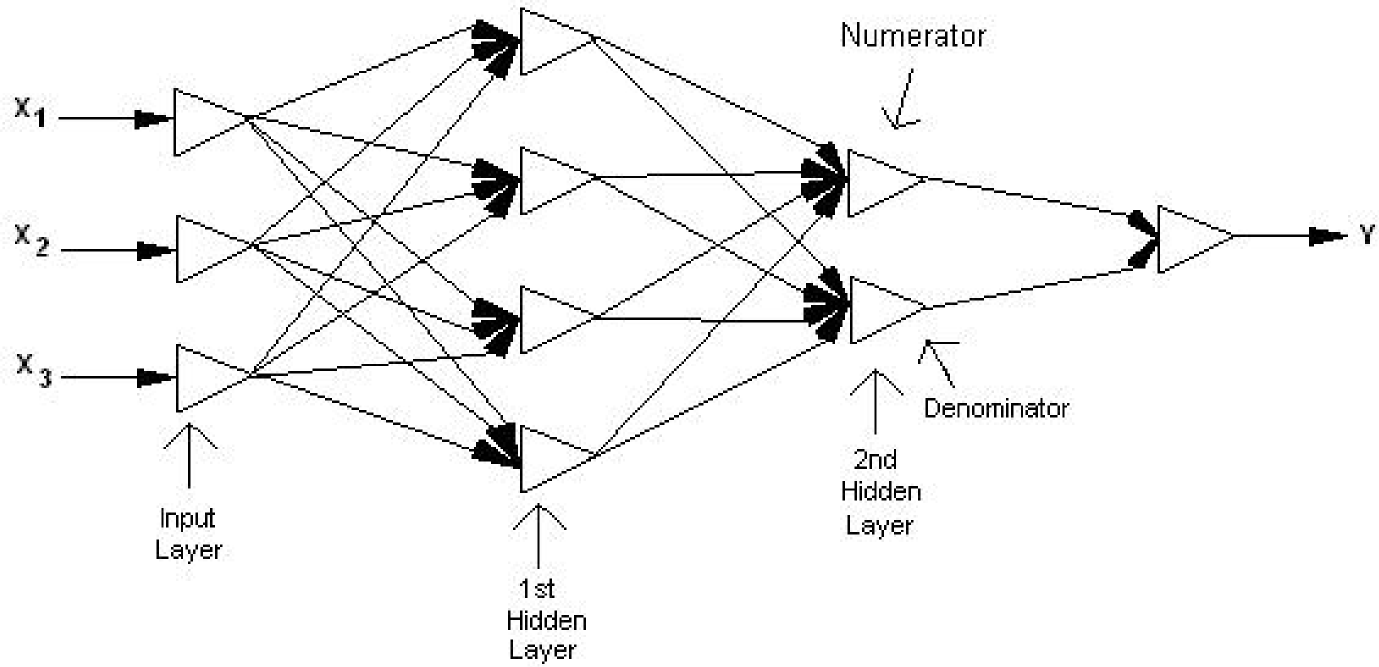
- **P(x,y) is the joint density of (X,Y). It can be shown that the above can be estimated from data as**

$$\hat{m}(x) = \frac{\frac{1}{mh}\sum_{i=1}^{m} K\left(\frac{x - X_i}{h}\right)Y_i}{\frac{1}{mh}\sum_{i=1}^{m} K\left(\frac{x - X_i}{h}\right)}$$

Function Approximation…8

- **This is known as the Nadaray and Watson Estimator (NWE).**
- **It is the basis of the Generalized Regression Neural Network. (GRNN).**
- **One major advantage of this NN is that estimates of Y are obtained directly without any training.**
- **We have Chosen to Study this Network and find ways of incorporating the Desirable Properties in our Data Reduction Model- NPCA-NN**

Schematic of GRNN



Further Mathematical Analysis of GRNN

- **In Classical Parametric Statistics, L_2 is also termed the Mean Squared Error (MSE).**

$$MSE(\hat{m}) = \mathrm{E}(\hat{m} - m)^2$$

- **Which can be decomposed into Variance and squared Bias**

$$MSE(\hat{m}) = Var(\hat{m}) + \left[\mathrm{E}\hat{m} - m \right]^2$$

Further Mathematical Analysis of GRNN..2

- **With a Change of Variable Let**

$$K_h(x, X_i) = \frac{1}{h} K\left(\frac{x - X_i}{h} \right)$$

- **To Compute** $MSE(\hat{m})$ **we need the mean and variance of**

$$\hat{m}$$

Further Mathematical Analysis of GRNN..3

- The (NW)-Estimator is once again

$$\hat{m}(x) = \frac{\dfrac{1}{mh}\displaystyle\sum_{i=1}^{m} K\left(\dfrac{x - X_i}{h}\right) Y_i}{\dfrac{1}{mh}\displaystyle\sum_{i=1}^{m} K\left(\dfrac{x - X_i}{h}\right)}$$

- Note that the denominator is the density estimate of P(x)

Further Mathematical Analysis of GRNN..4

- **ie**

$$\hat{P}(x) = \frac{1}{mh}\sum_{i=1}^{m} K\left(\frac{x - X_i}{h}\right) = \frac{1}{m}\sum_{i=1}^{m} K_h(x - X_i)$$

- **The (NW)-Estimator is Linear and can be written as:**

$$\hat{m}(x) = \sum_{i=1}^{m} v_i Y_i$$

Further Mathematical Analysis of GRNN..5

- **Therefore Means and Variance is easily computed**

$$E\left[\hat{m}(x)\right] = \sum_{i=1}^{m} v_i m(x_i)$$

$$Var\left[\hat{m}(x)\right] = \left(\sum_{i=1}^{m} v_i^2\right)\sigma^2$$

- **ie you are smoothing the points (x_i,m(x_i)) rather than (X_i,Y_i)**

Further Mathematical Analysis of GRNN..6

- **In Computing Means and Variance of NW-Estimator, Denominator and Numerator are treated separately ie for denominator:**

$$E\left[\hat{p}(x)\right] = EK_h(x, X)$$

$$Var\left[\hat{p}(x)\right] = \frac{1}{m} VarK_h(x, X)$$

- **Variance can be written as: variance = 2nd moments – (mean)^2**
- **To proceed, convolution is used followed by Taylor series expansion. Derivation is long**

Further Mathematical Analysis of GRNN..7

- **Final Theorem is:**
- **Bias Term: Let $u=x-X_i$**

$$\left(m''(x) + \frac{2m'(x)P'(x)}{P(x)} \right) \frac{h^2}{2} \int u^2 K_h(u)\,du$$

- **Variance Term**

$$\frac{\sigma^2(x)}{P(x)mh} \int K_h^2(u)\,du$$

Further Mathematic Analysis of GRNN…8

- These analysis answer two important concepts: Consistency and Rate of Convergence of NW-Estimator.
- In addition, In CFD of Reactive Flows, we want to look at the Bias terms.
- How does the Bias term depend on the Pattern of design points P(x) asymptotically?

Further Mathematatic Analysis of GRNN…9

- **Can you choose weight function K(u) such that the Bias is independent of P(x) ? Answer Yes-> Grid Free Solution of CFD using Neural Network**
- **An Estimator is said to be Consistent if as the sample size grows:**

$$\lim_{m\to\infty}\int\left[\hat{m}(x)-m(x)\right]^2 P(x)dx=0$$

- **Consistency does not tell us How Fast L_2 approaches 0. Here we look at the expectation of L_2 error.**

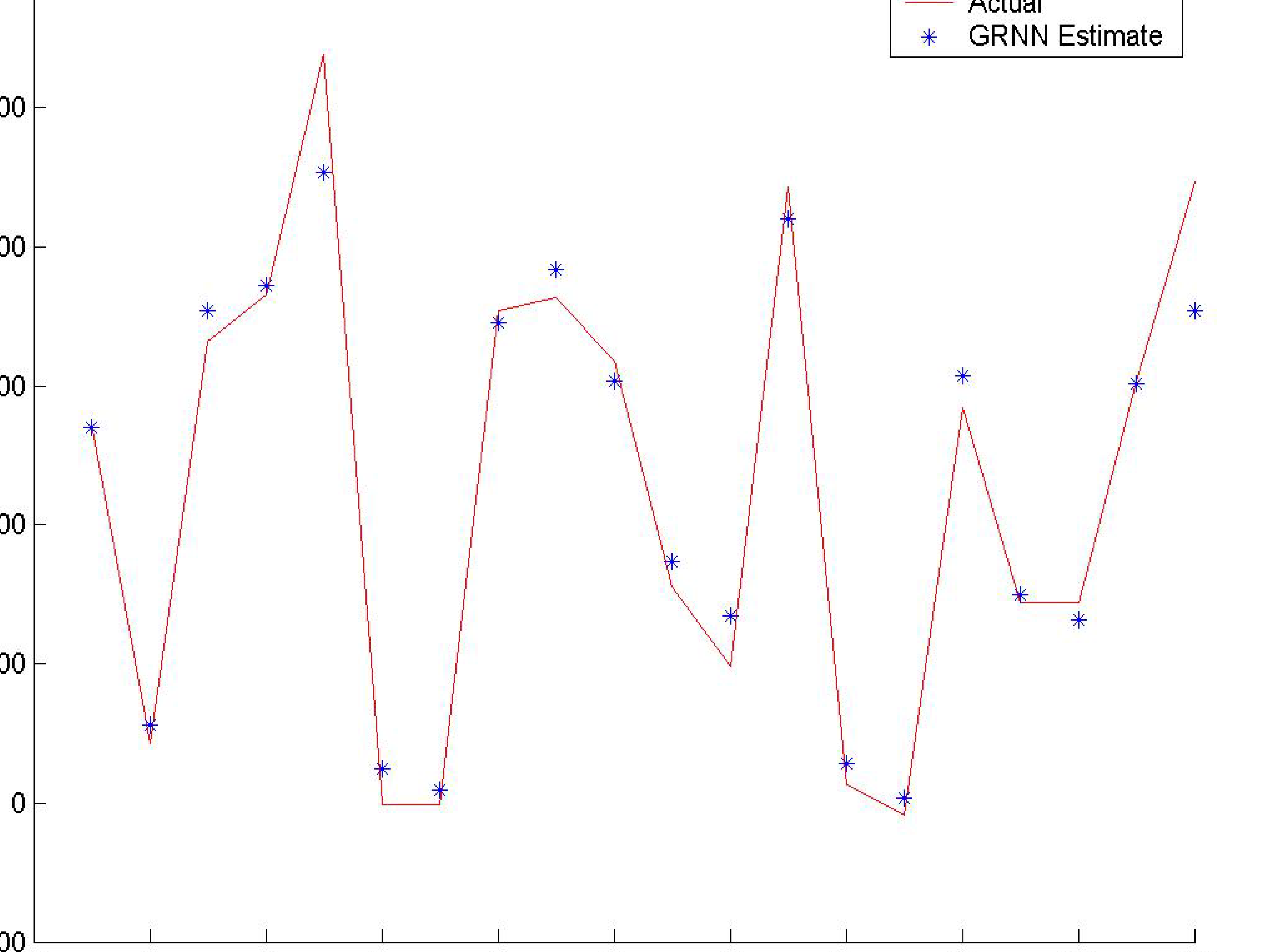
$$E\int\left|\hat{m}(x)-m(x)\right|^2 P(x)dx$$

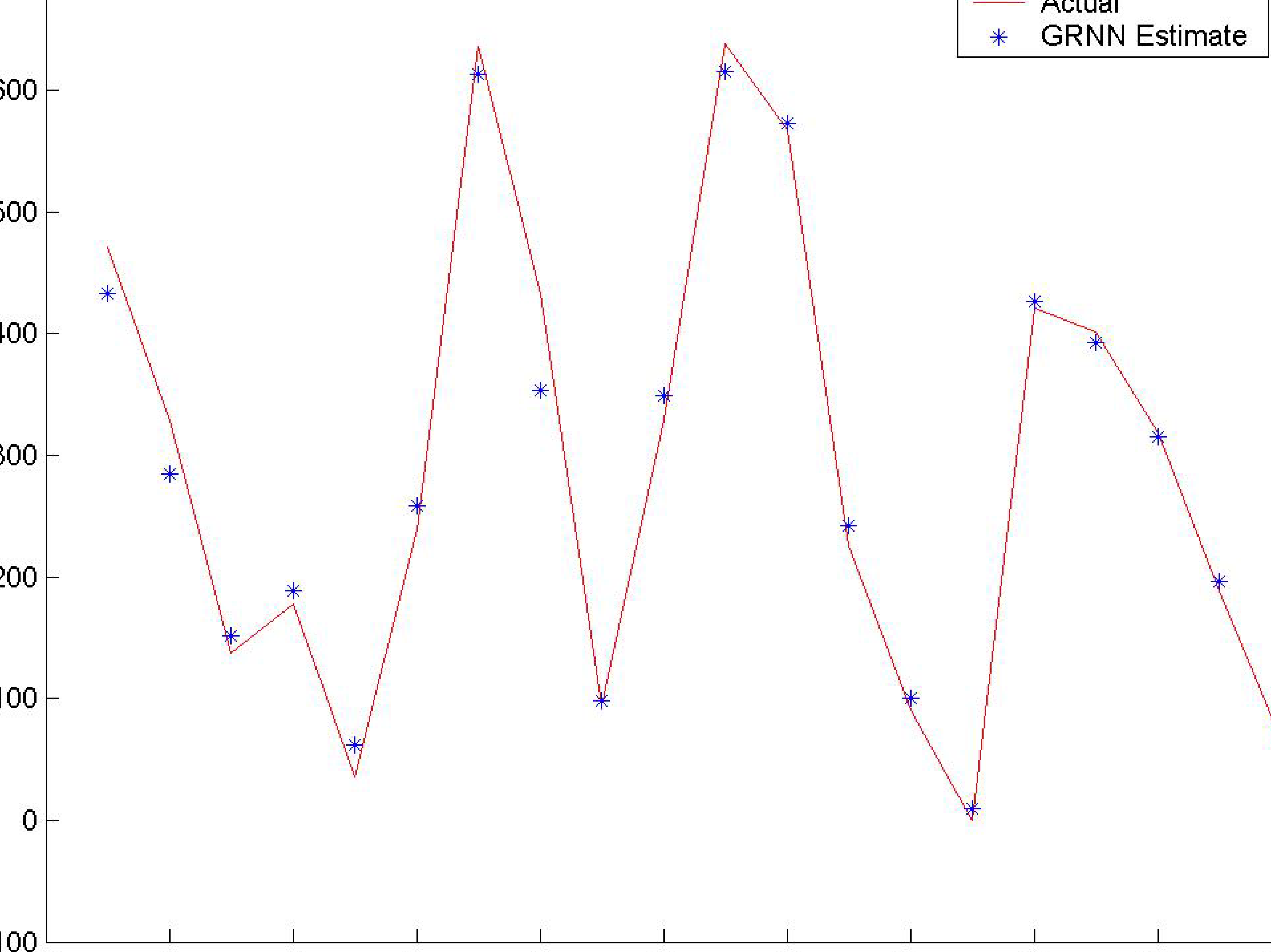
- **Difficult to analyze without imposing some smoothens assumptions on m(x)**

GRNN Results

- **3-Dimensional Test Function**

$$f(x_1, x_2, x_3) = 4x_1^2 + 2x_2^2 + x_3^2 - x_1 x_2 + x_2 x_3 - 5x_1 - 9x_2 + x_3$$


Actual
GRNN Estimate
00
00
00
00
00
0
00



Actual
GRNN Estimate
600
500
400
300
200
100
0
100

The Euler Solver

Overview

- The Euler Solver
- General Layout of the Program
- Results

The Euler Solver

- **The general incompressible flow of the governing equation:**

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \qquad (1.0a)$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+w\frac{\partial u}{\partial z}=-\frac{1}{\rho}\frac{\partial \rho}{\partial x}+\nu\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}\right) \qquad (1.0b)$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+w\frac{\partial v}{\partial z}=-\frac{1}{\rho}\frac{\partial \rho}{\partial y}+\nu\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}+\frac{\partial^2 v}{\partial z^2}\right) \qquad (1.0c)$$

$$\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial x}+v\frac{\partial w}{\partial y}+w\frac{\partial w}{\partial z}=-\frac{1}{\rho}\frac{\partial \rho}{\partial z}+\nu\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}+\frac{\partial^2 w}{\partial z^2}\right) \qquad (1.0d)$$

The 2D Euler Equation

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \qquad (1.1a)$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = 0 \qquad (1.1b)$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} = 0 \qquad (1.1c)$$

Euler equations in conserved form

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=0 \qquad (1.2)$$

$$U=\begin{bmatrix}\rho \\ \rho u \\ \rho v \\ \rho e\end{bmatrix}; \quad F=\begin{bmatrix}\rho u \\ \rho u^2+p \\ \rho uv \\ \rho uh\end{bmatrix}; \quad G=\begin{bmatrix}\rho v \\ \rho uv \\ \rho v^2+p \\ \rho vh\end{bmatrix} \qquad (1.3)$$

General Layout of the program

- **Program requirements**
- **Boundary and Initial Conditions**
- **Solution procedure**
- **Initializing the variables**

Program Requirements

- The grids should be generated independently and read in as part of the initial data.
- The initial values must be made available to the program and part of the initial data as well.
- The time-steps should be chosen according to CFD criteria for stability.

Boundary and Initial Conditions

- **The boundary conditions are calculated using extrapolation, interpolation and in some cases algebraic formulae.**
- **Since the Euler solver uses the time marching approach, it requires initial values of temperature, pressure etc.**

Solution Procedure

- **Step one**
  - **Set up the grid and the initial guess**
    - **These information are read into the program.**
    - **Apart form the grid and initial guess various other data must be fed in so that the problem is well defined**
      - **Boundary conditions**
      - **Coefficients of the Runge-Katta scheme**
      - **Relaxation parameter**
      - **Printing criteria etc**

Solution Procedure II

- **Step two**
  - **Computation of all relevant quantities**
    - **Cell areas and projective lengths along the two Cartesian axes**
    - **Radii of curvature at points of the blade surfaces.**

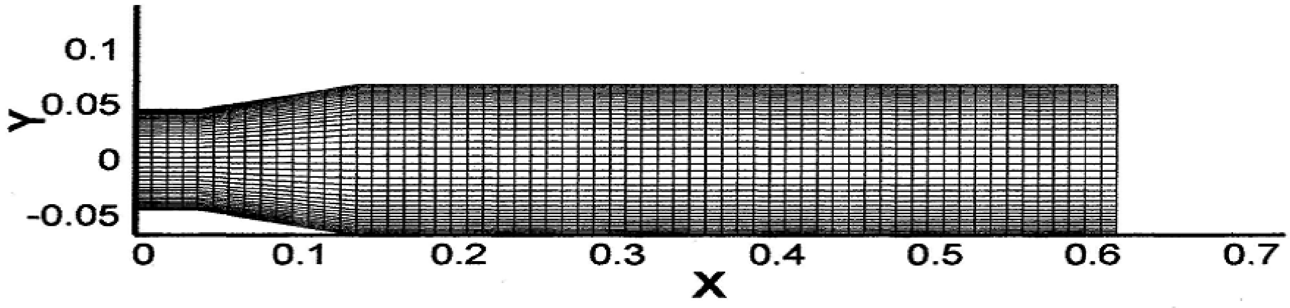
Initializing the variables

The variables to initialize are the ones that will be required in the approximating of the continuity, momentum and energy.

Problem

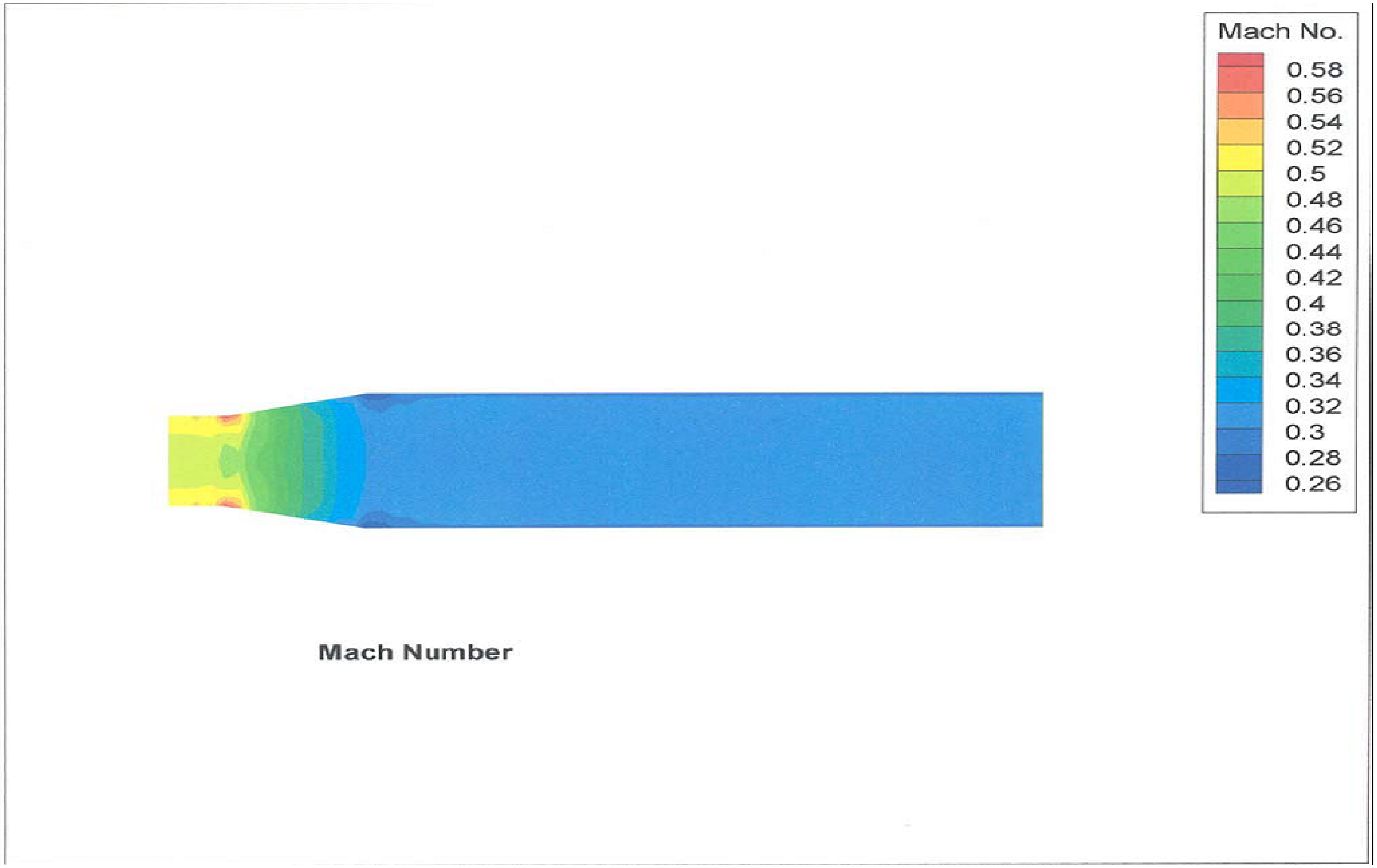
- **The code was tested on a simple problem, which was to solve for steady flow of air through a plane symmetrical diffuser. In this problem the viscosity and density of air were taken to be 1.91 X 10 –5 kg/ms and 1.21 kg/m^3, respectively. The flow profile at the inlet was assumed to be flat with a bulk velocity of 160 m/s. The flow at the outlet was assumed to be fully developed..**

Plane Symmetrical Diffuser

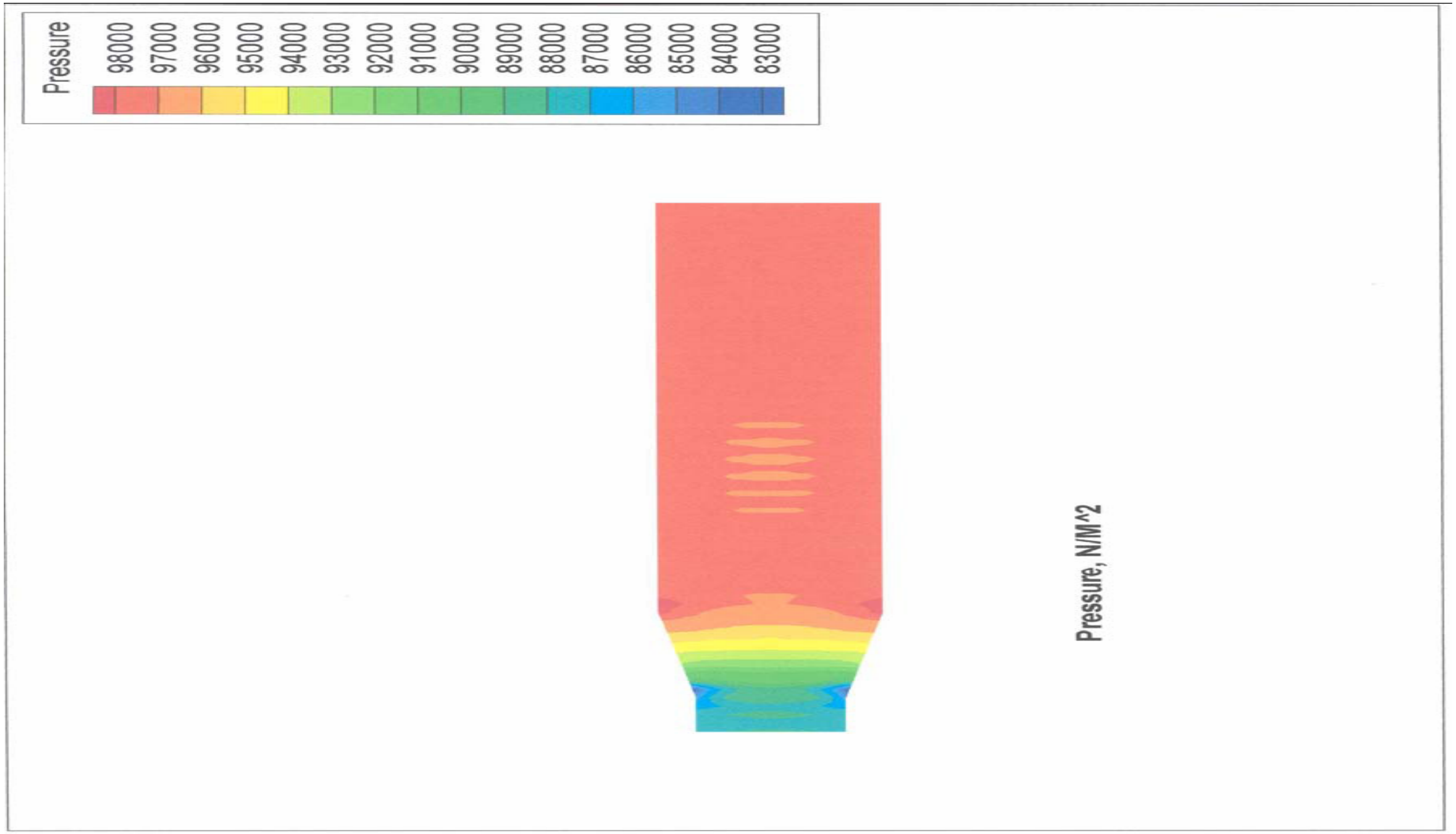


Grid Dimensions in M

Results for Mach Number



Results Static Pressure Distribution



Conclusion

- **Demonstrated the NPCA reduction technique for mechanism reduction on sample reaction mechanism**
- **Details of Mathematical Analysis Presented**
- **Description of Test Flow Solver**
- **Work is in progress**
- **Acknowledge the support of DOE Under grant numbers:**
  - **DE-FG26-00NT-40830**
  - **DE-FG26-03NT-41913**